

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ : C12Q 1/68, C12P 19/34 C07H 21/04	A1	(11) International Publication Number: WO 94/11532 (43) International Publication Date: 26 May 1994 (26.05.94)
(21) International Application Number: PCT/US93/10964 (22) International Filing Date: 12 November 1993 (12.11.93) (30) Priority data: 07/977,284 13 November 1992 (13.11.92) US (60) Parent Application or Grant (63) Related by Continuation US 07/977,284 (CIP) Filed on 13 November 1992 (13.11.92) (71) Applicant (for all designated States except US): THOMAS JEFFERSON UNIVERSITY [US/US]; 11th and Walnut Streets, Philadelphia, PA 19107 (US).		(72) Inventors; and (75) Inventors/Applicants (for US only) : PROCKOP, Darwin, J. [US/US]; 291 Locust Street, Philadelphia, PA 19108 (US). ALA-KOKKO, Leena [FI/US]; 317A High Road, Salem Harbour, Andalusia, PA 19020 (US). WILLIAMS, Charlene, J. [US/US]; 46 Long Bow Drive, Sewell, NJ 08080 (US). RITVANIEMI, Pertti [FI/FI]; Torppankatu 21 A 8, FIN-90230 Oulu (FI). BALDWIN, Clinton [US/US]; 37 Trowbridge Street, Cambridge, MA 02138 (US). HOPKINSON, Ian [GB/GB]; 9 Vennwood Close, Wenvoe, Cardiff (GB). AHMAD, Nilofer, Nina [US/US]; 432 East Allens Lane, Philadelphia, PA 19119 (US). (74) Agents: JOHNSON, Philip, S. et al.; Woodcock Washburn Kurtz Mackiewicz & Norris, One Liberty Place, 46th Floor, Philadelphia, PA 19103 (US). (81) Designated States: CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i>
(54) Title: METHODS OF DETECTING A GENETIC PREDISPOSITION FOR OSTEOARTHRITIS		
(57) Abstract The invention provides probes and primers for amplifying certain regions of genes for structural proteins of cartilage and methods for detecting mutations in these genes isolated from the nucleic acid of cells suspected of exhibiting mutant structural protein gene expression or having mutant structural protein genes. The invention also provides methods for determining a genetic predisposition for a disease that alters the structure or function of cartilage because of a mutation in a gene for a structural protein of cartilage in a mammal.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

**METHODS OF DETECTING A GENETIC
PREDISPOSITION FOR OSTEOARTHRITIS**

ACKNOWLEDGEMENT OF GOVERNMENT RIGHTS

Research for this invention was supported in part by
5 the National Institutes of Health Grants AR-381 88 and
AR-39740. The government may have certain rights in the
invention.

FIELD OF INVENTION

The present invention relates to the field of
10 compounds and methods for detecting genetic diseases linked to
anomalies of genes for collagens and other structural proteins
found in cartilage and joints.

BACKGROUND OF THE INVENTION

Osteoarthritis is a progressive disease of joints that
15 is a cause of serious disability in large numbers of people.
The disease is defined as a progressive degeneration of the
cartilaginous surfaces of joints that leads to stiffness, pain,
and loss of mobility. Degeneration of the cartilaginous
surface of joints seen in osteoarthritis can have a number of
20 causes. For example, severe trauma to a joint or a bacterial
infection in a joint can produce degeneration of the joint that
is either immediate or slowly progressive over many years. A
number of metabolic disturbances are also known to produce
degeneration of joints.

25 Cartilage and membranes that line joints are complex
structures. A major source of the strength of cartilage is the
fibrils of type II collagen. The fibrils of type II collagen

- 2 -

are stretched into three-dimensional arcades primarily by the presence of another group of macromolecules called proteoglycans. Proteoglycans are highly charged and, therefore, absorb water and salts and thereby extend the
5 arcades of type II collagen fibrils. As a result, a highly resilient structure is formed that can withstand the intermittent pounding and pressures that joints must undergo. In addition to proteoglycans and type II collagen, cartilage is known to contain at least four other kinds of collagens (types
10 VI, IX, X and XI) in lesser amounts than type II collagen. It is very likely that additional collagens will be discovered in cartilage in the future. In addition, it is clear that the matrix of cartilage also contains a number of other proteins that are still poorly characterized and that may contribute to
15 the structure and function of the tissue.

Collagens, proteoglycans and other proteins found in the matrix of cartilage are synthesized by cells embedded within the matrix. The matrix is actively synthesized during embryonic development of certain tissues and during periods of
20 growth. The rates of synthesis and degradation of the matrix are less during adult life. However, throughout life, a continual slow synthesis and degradation of cartilage occurs, particularly in response to the pressures associated with physical activity.

25 Cartilage itself has several different functions in the body. During embryonic development, transient tracks of type II collagen and probably other components of cartilage are formed in many structures. The tracks appear to serve as a guide for cell migration and a template for formatting of
30 skeleton and associated structures. In addition, cartilage serves as a precursor structure for many bones. During the development of long bones such as those of the arms and legs, cartilage is part of the growth plate in which cell growth occurs. More specifically, the cartilage grows away from the
35 midpoint of the long bone and is continually degraded and gradually replaced by bone itself. An additional function of cartilage is to give shape and form to tissues such as the nose

- 3 -

and ears. Many of the macromolecules found in cartilage are also present in the vitreous gel of the eye and account for the high viscosity of the vitreous. Still another major function of cartilage is to provide strength and resilience to
5 structures such as the intervertebral disc of the spine. In joints, it provides not only strength and resilience, but also the smooth surfaces for motion under heavy loads.

The degeneration of joint cartilages that occurs in osteoarthritis is caused by a failure of the cartilage to
10 maintain its structural integrity. In this process, the cartilage surface is eroded by physical pressures and is not adequately replaced by the new synthesis of cartilage. Instead of adequate repair of cartilage, secondary changes occur in the joint surface and in the joint. These changes include, for
15 example, inflammatory responses characterized by invasion of white cells and macrophages, abnormal deposition of mineral in the form of calcium and phosphate within the joint space and in the cartilage itself, deposition of fibers of type I and other collagens that are not normally part of cartilage or the joint,
20 abnormal growth of cartilage cells and matrix at locations adjacent to the joint surface and abnormal calcifications of the joints and associated structures. As part of the complex changes that occur, the cells of the cartilage or the invading cells from the blood stream begin to secrete degradative
25 enzymes that further contribute to the degradation of the joint structures.

In the more severe diseases of cartilage known as chondrodysplasias, serious defects in the formation of cartilage are apparent early in life and there is a failure of
30 joints to develop their normal size and shape. There is also a secondary failure of bone growth seen in these diseases. Moreover, there can be a failure of normal development of many tissues such as failure to achieve closure of normal partitions between oral and nasal passages, known as cleft palate, and
35 improper development of the vitreous gel of the eye that causes severe myopia and retinal detachment.

- 4 -

Research has demonstrated that some forms of osteoarthritis and related conditions are caused by mutations in the genes that code for and, therefore, determine the structure of the collagens that are the major source of the strength of cartilage. Mutations of collagen that have been defined include, for example, mutations in the gene for type II collagen and its precursor type II procollagen. These mutations are of two general kinds. One kind of mutation decreases the synthesis of type II procollagen. The second kind of mutation leads to the synthesis of a defective form of type II procollagen. As a result of these mutations, there is either a decrease in the normal level of type II collagen in cartilage and other tissues that contain the protein or the formation of abnormal type II collagen fibrils that do not have the strength of normal type II collagen fibrils and, therefore, cause the cartilage of joints to be degraded by normal wear and tear. The two kinds of mutations can also produce drastic effects during normal growth and development. As a result, some individuals who inherit some of the mutated genes develop severe chondrodysplasias and die in utero or shortly after birth. Alternatively, such individuals can have serious deformities such as dwarfism which shows severe malformation of joints and may be associated with conditions of severe myopia, myopia with retinal detachment and blindness, cataracts, cleft palate, and unusual facial appearance. Other similar mutations in the same genes may produce much milder effects and cause progressive generalized osteoarthritis in which affected individuals are apparently normal until middle age when they develop progressive stiffness, pain and then immobility of many joints. Mutations of the gene for type II procollagen and collagen have been shown to cause these disorders. Research suggests that some of the conditions are caused by similar mutations in other genes that code for other structural macromolecules found in cartilage which contribute to its normal resistance to wear and tear. Methods and compounds are therefore desired for the analysis and detection of mutations in genes, both for type II procollagen and for a series of

- 5 -

known collagens that are components of cartilage (types VI, IX, X and XI), as well as for still undiscovered collagens and other structural proteins that contribute to the normal strength and function of cartilage.

5 SUMMARY OF THE INVENTION

The invention provides probes and primers complementary to certain regions of collagen genes. Also provided are methods using such primers and probes for detecting mutations in a collagen gene sequence isolated from
10 cells from individuals suspected of exhibiting mutant collagen gene expression or containing mutant collagen genes. Also included are methods to detect mutations in the collagen genes of members of the individual's family.

More specifically, the invention provides a method
15 whereby an individual who has developed osteoarthritis or a related condition, or is suspected of developing osteoarthritis or a related condition, is tested to see if the individual has a mutation in the DNA of a gene for a structural protein that is a normal component of cartilage. This method provides that
20 the DNA sequence of a collagen gene is examined in an individual with osteoarthritis or a related condition. The DNA sequence is also compared with corresponding regions of a standard DNA from a series of individuals known not to have the disease in question. Any difference in the base sequence from
25 the DNA of the individual tested as compared to the standard sequence is then evaluated in terms of whether or not it indicates an increased likelihood of the individual suffering from osteoarthritis or related disorder. For the first member of a family tested, all or a substantial portion of DNA coding
30 for the gene is sequenced and compared to the standard sequence.

The invention also provides that once the location of the gene mutation causing the disease is known, it can be looked for in members of the first individual's family. For
35 each genetically predisposed individual family member, the mutation in the gene is expected to appear in the same position

- 6 -

in the collagen gene tested. A difference in base sequence of the DNA from the individual tested as compared with the standard sequence can be evaluated to determine how the mutation will effect expression of the gene. The potential for
5 the difference in DNA sequence to produce the disease can also be evaluated in terms of whether it changes an amino acid sequence that is critical for the normal functioning of the protein.

Methods further provide that DNA derived from the
10 cells of the test sample is analyzed to determine whether or not the collagen gene contains a mutation. If a mutation is found in the gene, a rapid test can be devised for other members of the patient's family to determine whether or not they have the same mutation.

15 The methods of the invention are particularly useful in detecting mutations in human collagen genes. It is believed that these methods will also be useful in detecting mutant collagen genes in mammals.

Intronic sequences provided by the invention are
20 useful for developing oligonucleotide primers to amplify and sequence genomic DNA from patients suspected of having mutations in the gene for the $\text{pro}\alpha\text{I (II)}$ chain of type II procollagen which cause various disorders.

Methods of the present invention for detecting
25 mutations in the gene for type II procollagen can readily be applied to detection of mutations in genes coding for other structural proteins found in cartilage and associated tissues. For example, the nucleotide sequences of these genes can be used to design oligonucleotide primers to amplify genomic DNA
30 or cDNA for the gene. The products obtained using PCR can then be used to define the base sequences of genomic DNA or cDNA. Mutations in the genes for these other collagens and structural proteins in matrix that cause osteoarthritis and related conditions can be detected in the same manner as mutations in
35 the gene for type II procollagen.

- 7 -

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a gene map of type II procollagen (COL2A1) with introns and exons designated by lines and crosshatched boxes respectively.

5 Figure 2 shows a schematic for PCR amplification of COL2A1 exon 9 including the region of the gene comprising polymorphism 225 and 252 and agarose gels containing the amplification products.

10 Figure 3 shows a schematic for PCR amplification of COL2A1 exons 27 and 28 including the region of the gene comprising an intron 26 polymorphism and agarose gels containing the amplification products.

DETAILED DESCRIPTION OF THE INVENTION

15 The present invention concerns methods of diagnosing specific kinds of osteoarthritis in which there is systemic degeneration of many joints without any apparent external cause. In particular, the invention relates to progressive generalized osteoarthritis that produces degeneration of the cartilage of many joints. It also concerns diseases in which
20 there is malformation of joints that is apparent at birth or during childhood and that leads to progressive degeneration of joints, as well as other symptoms such as dwarfism and severe malformation of the skeleton and other cartilaginous tissues. More particularly, the invention is directed to diagnosing
25 diseases characterized by degeneration of joints in which the presence of the disease in several members of the same family or similarity between the disease of a given patient and diseases seen in families indicates that the disease has a genetic origin. Therefore, the invention concerns not only
30 progressive generalized osteoarthritis, but also a group of related disorders that are generally defined as chondrodysplasias. Skeletal dysplasia or related disease involving abnormalities of growth and of cartilaginous structures and of tissues containing the same structural
35 proteins as cartilage including chondrodysplasias, epiphyseal dysplasia, metaphyseal dysplasia, spondyloepiphyseal dysplasia,

- 8 -

spondylometaphyseal dysplasia and arthro-ophthalmopathy (the Wagner-Stickler syndrome) may be diagnosed using the methods of the invention. Skeletal disorders such as scoliosis and related conditions involving abnormalities in the cartilage and
5 other components of the vertebrae column and back may also be diagnosed in accordance with the invention.

More specifically, the present invention provides methods for gene analysis whereby one can definitively establish the cause of osteoarthritis and of chondrodysplasias
10 in individuals and in members of the individuals' families. Such methods will make it possible to identify individuals in families who are predisposed to develop these diseases. In the case of severe chondrodysplasias that produce crippling deformities and that may even be lethal, such methods will be
15 useful for prenatal diagnosis. In the case of milder chondrodysplasias and osteoarthritis, the results of the gene analyses may be used to counsel individuals predisposed to develop the diseases concerning preventive exercise programs, life styles, choice of careers, and family planning. Also, the
20 information generated by the gene analyses using the methods of the present invention may be used to develop new rational therapies for diseases of collagen. Moreover, the information can be used to develop animal models for such diseases. For example, information may be useful to develop transgenic mice
25 that will provide new means of testing new agents to cure and prevent the diseases.

Methods of the invention provide that an individual who has developed osteoarthritis or a related condition or is suspected of developing osteoarthritis or a related condition,
30 is tested to see if a mutation in the DNA of a gene for a structural protein for example, type II procollagen, is present. The methods of the invention further provide that after a mutation causing the disease or diseases in one individual is found, it can be sought in members of the first
35 individual's family.

The gene COL2A1 which encodes the pro α I (II) chain of type II procollagen is provided and examined as a demonstration

- 9 -

of the methods of the present invention. This demonstration in no way limits the scope of the claimed invention. In the first stage of the methods of the invention, the DNA sequence of the COL2A1 gene is examined in an individual with a cartilage and joint disease such as osteoarthritis. This DNA sequence is compared with corresponding regions of a standard DNA from a series of individuals known not to have the disease in question. It is believed that this strategy will be useful for any disease of collagen which exhibits at least one mutation in a collagen gene of the diseased mammal. The DNA sequences of the genes tested can be genomic DNA or cDNA prepared from RNA derived from a sample of cells or tissues taken from the individual. DNA may also be extracted from bodily fluids containing lysed cells. The standard DNA sequence and structure of the COL2A1 gene can be obtained by reference to known sequences or those set forth herein in Figure 1 and listed in the literature or computerized data banks. Sequences of other collagen genes used in the methods of the present invention may also be obtained from databases or may be sequenced employing commonly used methods. See, for example, Sanger et al., DNA Sequencing With Chain-Terminating Inhibitors, *Proc. Natl. Acad. Sci. USA* 1977, 74, 5463-5467. Any difference in the base sequence from the DNA of the individual tested as compared to the standard sequence is then evaluated to determine whether it indicates an increased likelihood of the individual suffering from osteoarthritis or a related disorder. For the first member of a family tested, all or a substantial portion of DNA coding for the gene is sequenced and compared to the standard sequence. Sequencing of the first family member's DNA may be achieved by DNA sequencing techniques known in the art. These methods of the present invention are useful for all collagen genes.

As used herein, the term "collagen" includes procollagens and collagens types I to XVI, still undiscovered collagens similar to types I to XVI, and genes encoding proteins associated or comprising collagen polymers in tissue matrices.

- 10 -

The term "family member", as used herein, means individuals, including humans and other mammals, genetically related to one another in any degree, such as, for example, parent-child, siblings, cousins, etc. Such genetic relatedness
5 can be determined using standard methods known in the art including, for example, pedigree analysis or DNA "fingerprinting".

Moreover, the term "individual", as used herein, denotes a mammalian individual of any species, including
10 humans.

As illustrated by the demonstration using type II collagen, methods of detection of mutations in a collagen gene comprise several steps. One step involves selecting cells suspected of comprising a mutated collagen gene. Another step
15 includes isolating genomic DNAs from selected cells or preparing cDNAs from selected cells. After the genomic DNA or cDNA is isolated, larger amounts of the gene sequences of interest are prepared by amplifying the DNA with the polymerase chain reaction (PCR). The DNA produced by the PCR are then
20 analyzed for the presence of a disease-causing mutation. The preferred strategy of analysis is to first screen the PCR products with a relatively rapid technique such as denaturing gradient gel electrophoresis (DGGE) that enables one to decide whether or not a specific PCR product from one region of the
25 gene does or does not have a mutation such as a single base difference between the two alleles of the gene. PCR products detected as probably having a mutation by such a technique are then analyzed further by a technique such as dideoxynucleotide sequencing that provides the detailed base sequence that
30 defines the mutation. The methods comprise the aforementioned steps and further comprise comparing the sequence of the collagen gene or other genes for structural proteins of cartilage containing the mutation to corresponding regions of a family member's structural protein genes and determining if
35 the mutation is present in the family member's genes.

It is believed that all of the methods of the present invention are useful to detect any mutation in all procollagen

- 11 -

and collagen genes, including, for example, procollagens and collagens I to XVI, still undiscovered procollagens and collagens similar to types I to XVI, as well as other genes associated with collagen structure, such as proteoglycans. It is also believed that these methods are also useful in mammals other than humans.

A second step of the invention provides that once the location of the mutation in a gene causing the disease is known, it can be sought in members of the first individual's family. For each genetically predisposed individual family member, the mutation in the gene is expected to appear in the same position in the structural protein gene tested. For example, in Family A, the genetic mutation may be at position 30; and for Family B, the genetic mutation may be at position 505. In accordance with the methods of the invention, testing the family members can be done by comparing corresponding regions of family member's genes and determining if the mutation is present in the family member. This evaluation of a difference in base sequence of the DNA from the individual tested as compared with the standard sequence can be evaluated in terms of whether it is a disease-causing mutation by determining whether the mutation changes the level of expression of the gene in terms of the rate at which the gene is transcribed into RNA, the rate at which the initial RNA is processed into mRNA and the rate at which the mRNA can be effectively used to synthesize mature collagen matrices, such as pro α I (II) chains of type II procollagen. The potential for the difference in DNA sequence to produce the disease can also be evaluated in terms of whether it changes an amino acid sequence that is critical for the normal functioning of the protein by strategies used by those familiar in the art. The strategies include demonstration that the same gene mutation is not present in individuals unaffected by the disease in the same family of the general population, detailed linkage analysis of co-inheritance of the mutation with the disease phenotype in large families or a series of families with the same mutation, and structure-function studies on the mutated

- 12 -

protein obtained either by isolation of the protein from tissues of affected individuals or expression of the mutated gene in a recombinant system.

In accordance with the methods of the invention, DNA
5 is extracted from a test sample of cells of the family member to be tested by conventional techniques, such as lysis of the cells with sodium dodecyl sulfate (SDS) and digestion of protein with proteinase K, followed by extraction with phenol and chloroform, and ethanol precipitation as described by
10 Maniatis et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1982, pp 280-281. A sample of cells can be taken from any type of tissues; for example, a piece of skin, a sample of blood, or by scraping of the interior of the mouth. Alternatively, mRNA can
15 be extracted from the test sample and cDNA synthesized with reverse transcriptase, and the resulting cDNA used for analysis. Although cartilage cells cannot regularly be obtained from patients, many other cells including white blood cells have been shown to contain small amounts of the mRNAs for
20 protein synthesized by cartilage and the mRNAs can be analyzed after conversion to cDNAs as reported by Chan and Cole, *Journal of Biological Chemistry* 1991, 266, 12487.

Following extraction, DNA derived from the cells of the test sample is analyzed to determine whether the structural
25 protein gene contains a mutation. If a mutation is found in the gene, a rapid test can be devised for other members of the patient's family to determine whether they have the same mutation. DNA and cDNA from other structural protein genes can also be used in this method of the invention.

30 Although the methods of the invention have been demonstrated in the first instance in human beings, it is expected that they will be useful in other mammalian species, particularly commercially important species and in laboratory animals used as models of human disease. For example, it is
35 believed that these methods will be particularly useful in detecting mutant collagen genes and transcripts in transgenic animals comprising mutant collagen genes or transcripts.

- 13 -

Some of the sequences for the normal COL2A1 gene can be found in the literature. For example, in Strom and Upholt, *Nucl. Acids Res.* 1984, 12, 1025-1038; Cheah et al., *Proc. Natl. Acad. Sci. USA* 1985, 87, 2555-2559; Sangiorgi et al., *Nucl. Acids Res.* 1985, 13, 2207-2225; Nunez et al., *Gene* 1986, 44, 11-16; Su et al., *Genomics* 1989, 4, 483-441; Vikkula and Peltonen, *FEBS Lett.* 1989, 250, 171-174; Upholt, *Collagen Vol. 4*, CRC Press, Baton Rouge, FL, 1989, pp 31-49; Ala-Kokko and Prockop, *Genomics* 1990, 8, 454-460. However, the sequences for a major part of the introns used for the synthetic oligonucleotide primers employed for the analyses described here are provided as part of the present invention. Those skilled in the art recognize that rapid analysis of a gene by these procedures requires a series of oligonucleotides that are specifically targeted to regions of the gene under analysis. However, many base sequences are not efficient targets for oligonucleotide primers. In the case of procollagen and collagen genes, the number of efficient target sites is limited by the high GC content and repetitive nature of the coding sequences. Also, most of the genes contain a large number of introns, and many of the introns are too short to offer many potential target sites (see Figure 1). Therefore, the present invention includes efficient oligonucleotide primers for amplification of the type II procollagen gene by PCR, efficient oligonucleotide primers for analysis of PCR products of the type II procollagen gene by DGGE, and efficient oligonucleotide primers for dideoxynucleotide sequencing of PCR products of the type II procollagen. Table I presents specific oligonucleotide primers for amplification of the human type II procollagen gene by PCR and for dideoxynucleotide sequencing of the PCR products. Table II presents a series of neutral sequence variants detected in the type II procollagen gene that are important to avoid as target sites for primers, for defining haplotypes of the gene, and evaluating putative disease-causing mutations in the gene. Table III presents oligonucleotide primers that are efficient for amplifying regions of the gene

- 14 -

by PCR in a form that makes the PCR products suitable for screening for mutations by the technique of DGGE.

The analytical methods that are part of the invention involve the PCR. One skilled in the art would recognize that

5 there are many commonly employed schemes for amplifying nucleic acid sequences using PCR. See, for example, *PCR Protocols, A Guide to Methods and Applications*, Innis et al, Eds., Academic Press, New York, 1990, and *Current Communications, Polymerase Chain Reaction*, Ehrlich et al., Eds., Cold Spring Harbor

10 Laboratory Press, Cold Spring Harbor, NY, 1989. The PCR methods of the invention for amplifying nucleic containing at least one mutation in a collagen gene comprise several steps. The steps enumerated herein are merely exemplary of the steps commonly employed by one skilled in the art performing PCR.

15 Other steps may be added to achieve DNA amplification. Included in the invention are steps for selecting cells suspected of comprising a mutated collagen gene and isolating nucleic acid from the cells. Initial PCR steps comprise contacting the nucleic acid with a first primer and a second

20 primer, extending the first primer to create an extension product, contacting the extension product with a second primer, and extending the second primer to create another extension product. The extension products of the first round of synthesis may be longer than the extension products of the

25 second round of synthesis. Extension products of the first round of synthesis are contacted with the primers comprising complementary sequences. The primers used in this step may be the same primers used in the first round of synthesis. These primers are extended to form "secondary" extension products.

30 These "secondary" extension products may be shorter than the initial extension products from which they were synthesized. Amplification steps comprise amplifying the "secondary" extension products using PCR. The initial runoff DNA extension products will be diluted during amplification. The

35 amplification steps will favor synthesis of the "secondary" extension product DNAs having a length determined by the probes used. Once the PCR steps are completed, steps are provided for

- 15 -

detecting the presence or absence of said mutation in at least one extension product. It is believed that these methods will be useful for analyzing cartilage matrix protein genes, especially in humans. It is also believed that these methods
5 are useful to analyze the matrix protein genes of other proteins found in collagen matrices, such as non-collagenous structural protein of cartilage. Further, the PCR methods of the invention are useful for determining if a mammal has a genetic predisposition for a disease exhibiting a mutant
10 collagen gene. These methods include further steps of comparing the sequence of the collagen gene containing the mutation to corresponding regions of a family member's collagen genes and determining if the mutation is present in the family member's collagen genes. It is believed that these methods
15 will also be useful in other mammals.

Methods of the present invention for detecting mutations in the gene for type II procollagen can readily be applied to detection of mutations in genes coding for other structural proteins found in cartilage and associated tissues.
20 For example, the nucleotide sequences of the genes for types VI, IX, X and XI collagens can be used to design oligonucleotide primers to amplify genomic DNA or cDNA for the gene using PCR. The products obtained using PCR can then be used to define the base sequences of genomic DNA or cDNA.
25 Therefore, mutations in the genes for these other collagens and structural proteins in matrix that cause osteoarthritis and related conditions can be detected in the same manner as mutations in the gene for type II procollagen.

The following examples are illustrative of the
30 invention. It is understood that this invention is not limited by these illustrative examples but solely by the claims appended hereto.

EXAMPLES

Example 1

35 A series of procedures were developed for amplifying important regions of the gene for type II procollagen (COL2A1)

- 16 -

by the PCR so that adequate amounts of DNA were generated for analysis. A further series of procedures made it possible to directly analyze the DNA produced by the PCR and define its base sequences.

5 Figure 1 presents a diagram of the human COL2A1 gene indicating the location of the 54 exons of the gene. The figure also indicates, using short horizontal lines, the regions of the genes that were amplified with the use of appropriate oligonucleotide primers and conditioned for
10 amplification of the gene by PCR. Table I presents the sequencing and location of the oligonucleotide primers defined by the specific base sequences of the gene beginning with number 1 at the 5'-end of the gene. Ala-Kokko and Prockop, *Genomics* 1990, 8, 454-460, disclosed all the coding sequences
15 of the COL2A1 gene but no more than 40 intronic bases immediately flanking most of the introns.

Examples of primers effective for amplification of sequences by PCR and sequencing of the PCR products presented in Table I are within regions of the introns which have not
20 been disclosed previously. These sequences and primers are included in the present invention. The steps whereby genomic DNA from an individual from sources such as white blood cells or any other cells in the body can be used in the procedure to detect variations in sequences are as follows. The DNA
25 template is amplified using PCR. The PCR products can be loaded directly on an agarose gel and electrophoresed and analyzed or they can be diluted and subjected to a second round of PCR amplification. PCR products obtained from the first PCR reaction in which the regions of the gene symmetrically
30 amplified were examined by electrophoresis in an agarose gel stained with ethidium bromide. The conditions of the experiment were adequate to generate an intense single band of DNA, an observation indicating that the target region was selectively amplified. If a second PCR amplification is
35 carried out, the products are purified and sequenced. The second PCR in which the product of the first PCR is asymmetrically amplified generates a single-stranded DNA which

- 17 -

can be directly used for sequencing. The results indicate the presence of two major bands of DNA, one double-stranded DNA and the second single-stranded DNA appropriate for sequencing. Analysis of the base sequences of PCR products from the second asymmetric PCR were analyzed by dideoxynucleotide sequencing using methods known in the art. The sequencing autoradiograms were of high enough quality to be able to detect single-base variations in which one allele of the gene has one base and the same position of the second allele from the same gene has a different base. For example, patients 1 and 5 had an A in one allele and a T in the other allele at position +5 of intron 9 of the COL2A1 gene, whereas patients 2, 3, and 4 had only a T in this position. Moreover, patient 4 had an A and a G in position +45 of intron 9, whereas the other patients had only a G. Similarly, DNA from one patient showed both a C and a T at position -47 of intron 26 of the COL2A1 gene, whereas a different patient had a C at this position and the patient in the right-hand four lanes had a T in this position. These results demonstrate that the procedures outlined here are adequate to detect single-base variations in a single allele.

- 68 -

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19
 - (B) TYPE: NUCLEIC ACID
 - (C) STRANDEDNESS: SINGLE
 - (D) TOPOLOGY: LINEAR
- (iv) ANTI-SENSE: YES
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 93:
GACTCCAGGC TACCACGAA (19)
- (2) INFORMATION FOR SEQ ID NO: 94:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18
 - (B) TYPE: NUCLEIC ACID
 - (C) STRANDEDNESS: SINGLE
 - (D) TOPOLOGY: LINEAR
 - (iv) ANTI-SENSE: NO
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94:
CATGGAGGAG TGATATTC (18)
- (2) INFORMATION FOR SEQ ID NO: 95:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19
 - (B) TYPE: NUCLEIC ACID
 - (C) STRANDEDNESS: SINGLE
 - (D) TOPOLOGY: LINEAR
 - (iv) ANTI-SENSE: NO
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 95:
CTGCTGTGGA GAATTGTTT (19)
- (2) INFORMATION FOR SEQ ID NO: 96:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18
 - (B) TYPE: NUCLEIC ACID
 - (C) STRANDEDNESS: SINGLE
 - (D) TOPOLOGY: LINEAR
 - (iv) ANTI-SENSE: NO
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 96:
TGCTGTGGAG AATTGTTT (18)
- (2) INFORMATION FOR SEQ ID NO: 97:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20
 - (B) TYPE: NUCLEIC ACID
 - (C) STRANDEDNESS: SINGLE
 - (D) TOPOLOGY: LINEAR
 - (iv) ANTI-SENSE: NO
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 97:
CAATGCGGGC TGCCTCCTTG (20)
- (2) INFORMATION FOR SEQ ID NO: 98:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18
 - (B) TYPE: NUCLEIC ACID
 - (C) STRANDEDNESS: SINGLE
 - (D) TOPOLOGY: LINEAR
 - (iv) ANTI-SENSE: NO
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 98:
AATGCGGGCT GCCTCCTT (18)
- (2) INFORMATION FOR SEQ ID NO: 99:

- 103 -

CLAIMS

What is claimed is:

1. A method of determining a genetic predisposition for a disease caused by a mutation in the DNA of a gene encoding a structural protein that is a normal component of cartilage comprising:
 - selecting cells from an individual suspected of having a mutated structural protein gene;
 - isolating nucleic acid from the cells;
 - 10 amplifying regions of the isolated nucleic acid that contain the DNA of a specific structural protein gene using specifically targeted oligonucleotide primers;
 - analyzing the DNA of the structural protein gene for the presence of a mutation using the primers;
 - 15 detecting the presence of the mutation in at least one structural protein gene;
 - comparing the sequence of the gene containing the mutation to corresponding regions of a family member's collagen genes; and
 - 20 determining if the mutation is present in the family member's genes.
2. The method of claim 1 wherein the DNA analysis step comprises sequencing PCR products using the primers.
3. The method of claim 1 wherein the DNA analysis
25 step comprises analyzing PCR products by denaturing gradient gel electrophoresis.
4. The method of claim 1 wherein the amplification step comprises:
 - selecting cells suspected of having a mutated
30 structural protein gene;
 - isolating nucleic acid from the cells;
 - contacting the nucleic acid with at least two primers;
 - synthesizing primer extension products from the primers;

- 104 -

amplifying the extension products using polymerase chain reaction; and

detecting the presence of a mutation in at least one extension product.

5 5. The method of claim 1 wherein the structural protein gene is a collagen gene.

6. The method of claim 1 wherein the structural protein gene is a COL2A1 gene.

7. The method of claim 5 wherein the gene encodes
10 a non-collagenous structural protein of cartilage.

8. The method of claim 2 wherein at least one of the primers comprises a sequence identified in Table I.

9. The method of claim 3 wherein at least one of the primers comprises a sequence identified in Table III.

15 10. The method of claim 1 wherein the disease is osteoarthritis.

11. The method of claim 1 wherein the disease is a skeletal dysplasia or related disease involving abnormalities of growth and of cartilaginous structures and of tissues
20 containing the same structural proteins as cartilage, selected from the group consisting of chondrodysplasias, epiphyseal dysplasia, metaphyseal dysplasia, spondyloepiphyseal dysplasia, spondylometaphyseal dysplasia and arthro-ophthalmopathy.

12. The method of claim 1 wherein the disease is a
25 skeletal disorder involving abnormalities in the cartilage and other components of the vertebral column and back.

13. The method of claim 12 wherein the disease is scoliosis.

- 105 -

14. A primer comprising a sequence complementary to an intronic sequence from a collagen gene.

15. A primer of claim 14 wherein the intronic sequence is an intronic sequence from a mammalian procollagen
5 gene.

16. A primer of claim 15 wherein the procollagen gene is a COL2A1 gene.

17. A primer for use in detecting mutations in a mammalian gene for a structural protein of cartilage comprising
10 a sequence identified in Table III.

18. A primer for use in detecting mutations in a mammalian gene for a structural protein of cartilage comprising a sequence identified in Table I.

1 / 3

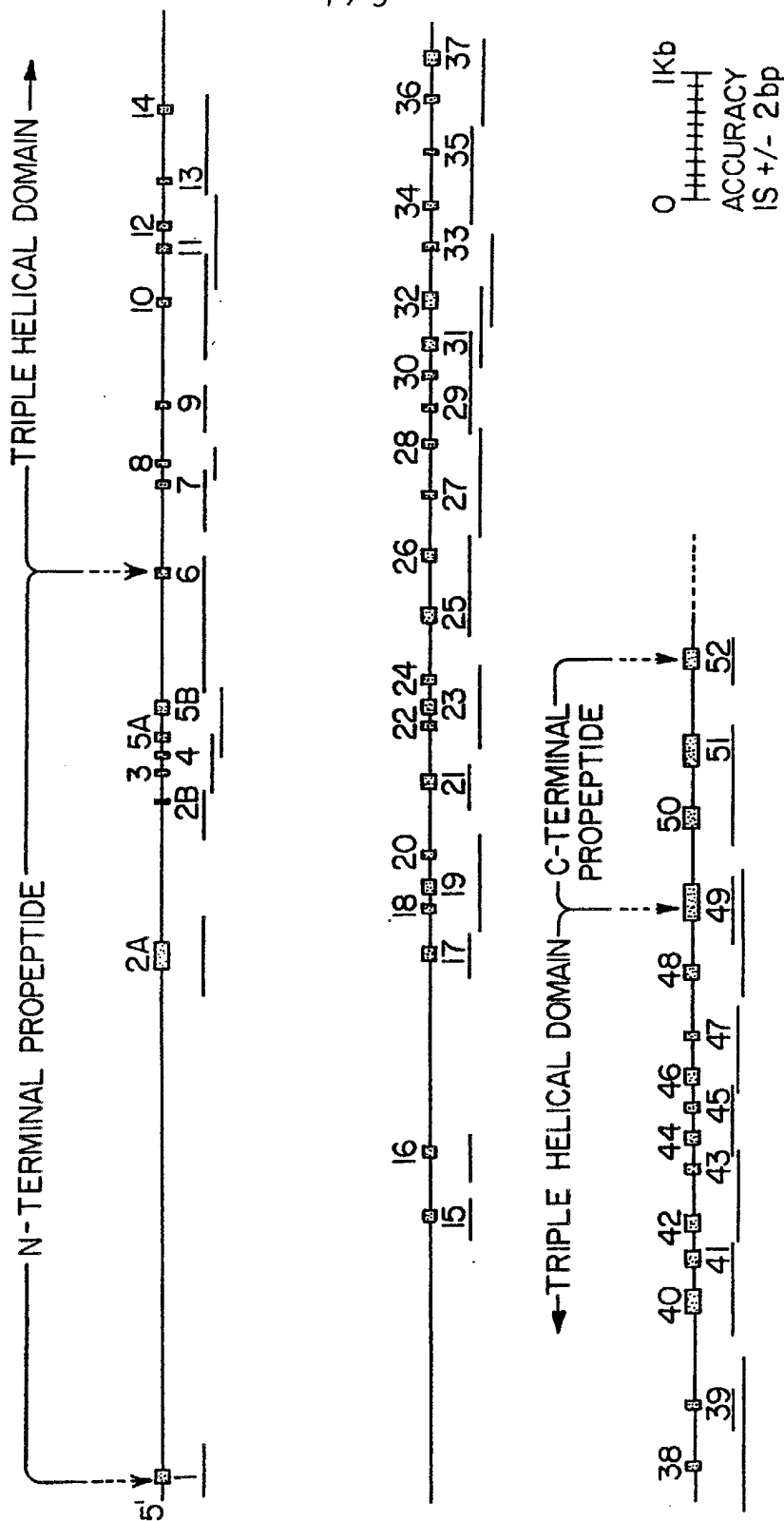
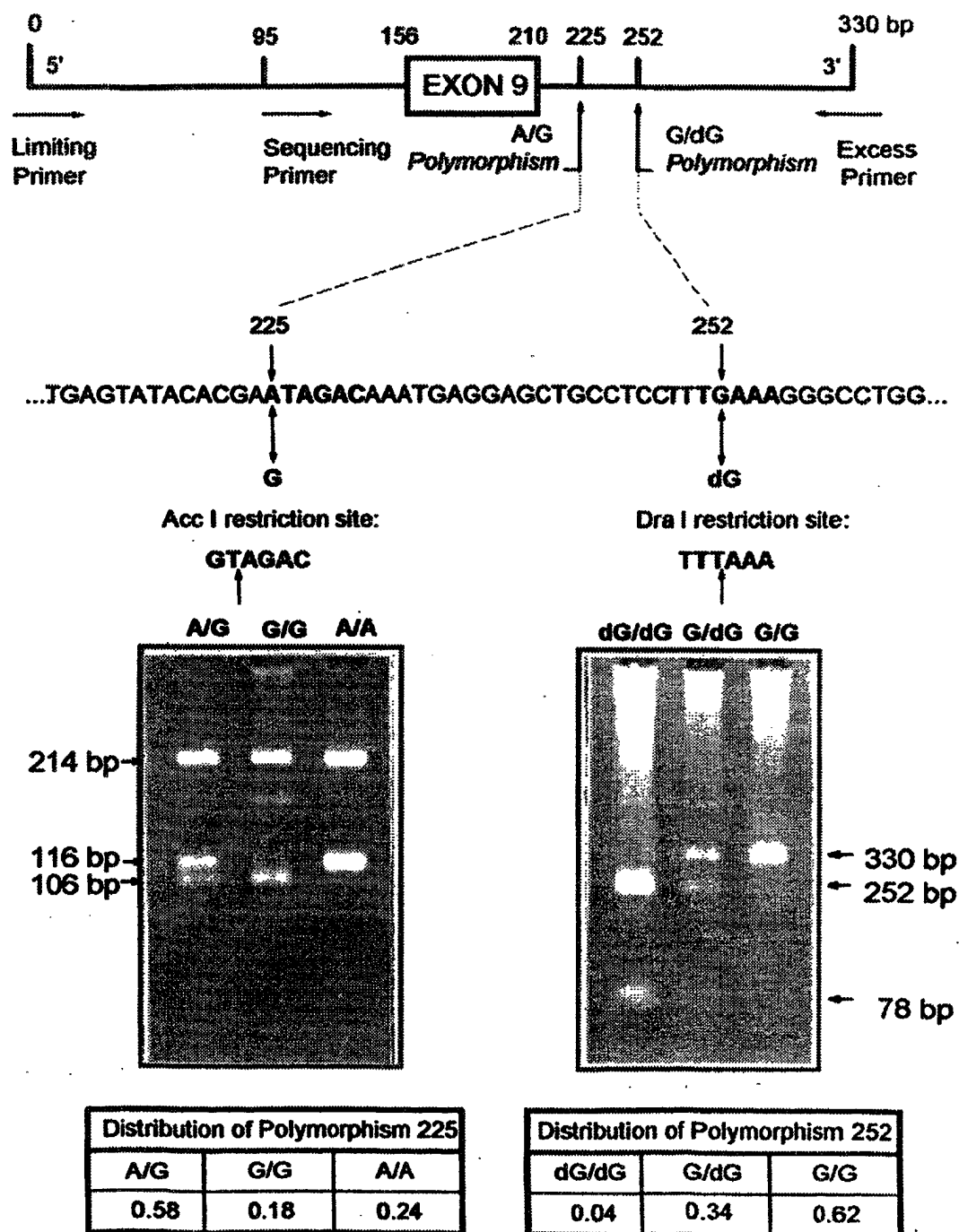
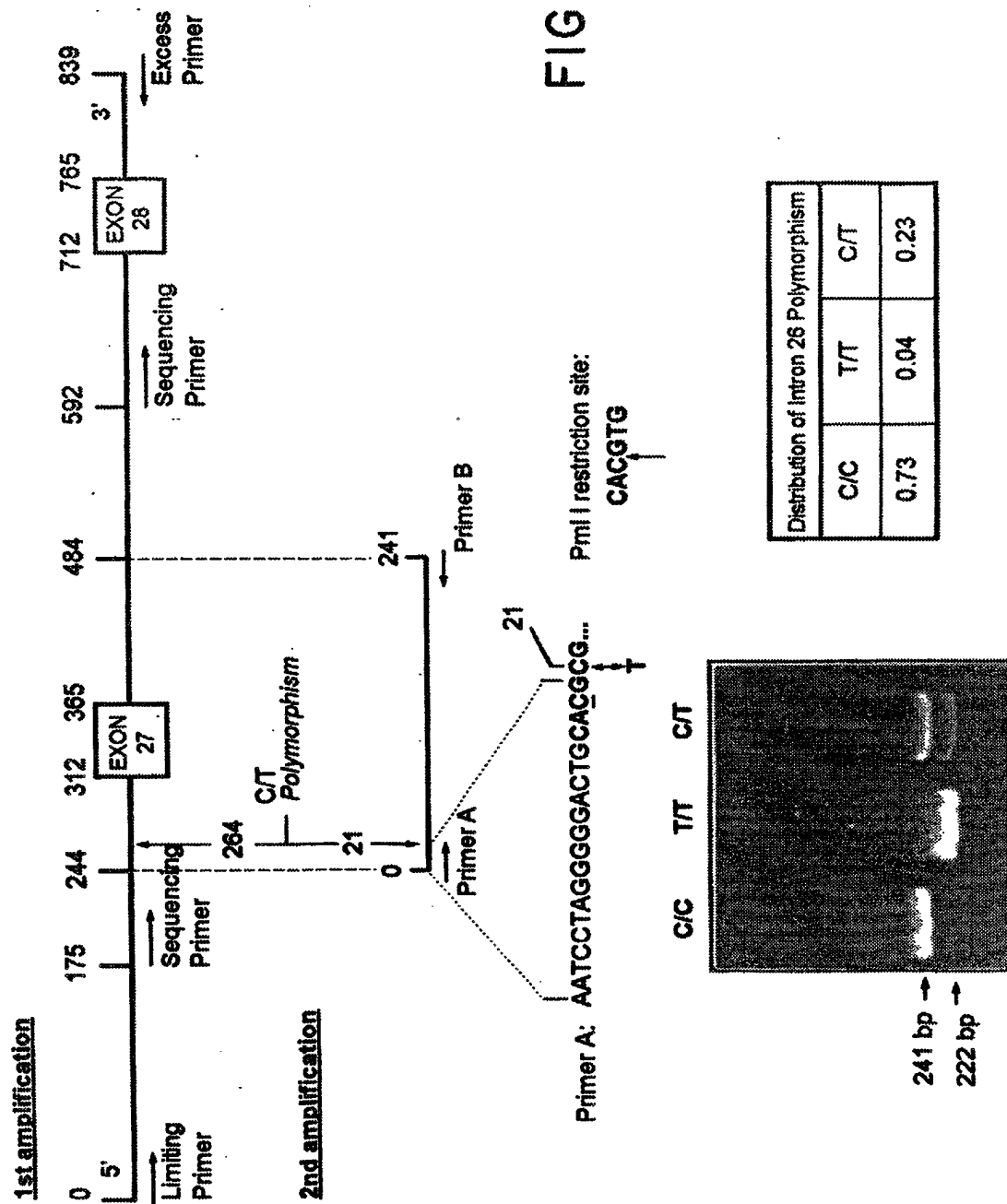


FIG. 1

FIG. 2^{2/3}

3/3

FIG. 3



INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/10964

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :C12Q 1/68; C12P 19/34; C07H 21/04

US CL :435/6, 91.1, 91.2; 536/24.33; 935/77,78

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 91.1, 91.2; 536/24.33; 935/77,78

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A 5,045,449 (ALA-KOKKO et al.) 03 SEPTEMBER 1991	1-18
Y	GENOMICS, Vol. 8, issued 1990, L. Ala-Kokko et al., "Completion of the Intron-Exon Structure of the Gene for Human Type II Procollagen (COL2A1): Variations in the Nucleotide Sequences of the Alleles from Three Chromosomes", pages 454-460, see pages 455-458, "Results".	1-18



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be part of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"A" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

18 FEBRUARY 1994

Date of mailing of the international search report

MAR 07 1994

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. NOT APPLICABLE

Authorized officer

STEPHANIE W. ZYTOMER, PH.D.

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/10964

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	THE JOURNAL OF BIOLOGICAL CHEMISTRY, Vol. 266, No. 19, issued 05 JULY 1991, D. Chan et al., "Low Basal Transcription of Genes for Tissue-Specific Collagens by Fibroblasts and Lymphoblastoid Cells", pages 12487-12494, see pages 12488-12491, "Amplification..."	1-18
Y	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCE, USA, Vol. 82, issued MAY 1985, Cheah et al., "Identification and Characterization of the Human Type II Collagen Gene (COL2A1)", pages 2555-2559, see pages 2557-2559, Figures 2-4	1-18
Y	NUCLEIC ACIDS RESEARCH, Vol. 13, No. 7, issued 1985, Sangiorgi et al., "Isolation and Partial Characterization of the Entire Human Proalpha1(II) Collagen Gene", pages 2207-2225, see pages 2214, 2216, 2218, 2219, Figures 5, 7, 9 and Table 1	1-18

INTERNATIONAL SEARCH REPORT

Int ional application No.

PCT/US93/10964

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

APS, BIOSIS, DERWENT BIOTECH ABS, CURRENT BIOTECH ABS, EMBASE, LEFE SCIENCES COLL, MEDLINE, WPI, DISS ABS ONLINE, CA SEARCH; Search terms: collagen, procollagen, mutat?, nucleotide sequence, structural, intron, detect, assay?, PCR, polymerase

L3 ANSWER 5 OF 5 CAPLUS COPYRIGHT 2003 ACS
 AN 1994:597648 CAPLUS
 DN 121:197648
 TI PCR-based methods of detecting a genetic predisposition for osteoarthritis
 IN Prockop, Darwin J.; Ala-kokko, Leena; Williams, Charlene J.; Ritvaniemi,
 Pertti; Baldwin, Clinton; Hopkinson, Ian; Ahmad, Nilofer Nina
 PA Thomas Jefferson University, USA
 SO PCT Int. Appl., 12 pp.
 CODEN: PIXXD2
 DT Patent
 LA English
 FAN.CNT 1

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	WO 9411532	A1	19940526	WO 1993-US10964	19931112
	W: CA, JP, US				
	RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
	US 5558988	A	19960924	US 1992-977284	19921113
	US 5948611	A	19990907	US 1995-256426	19950203
PRAI	US 1992-977284		19921113		
	WO 1993-US10964		19931112		

AB Probes and primers for amplifying certain regions of genes for structural proteins of cartilage and methods for detecting mutations in these genes isolated from the suspected cells are provided. The methods are useful for detg. a genetic predisposition. for a disease that alters the structure or function of cartilage because of a mutation in a gene for a structural protein of cartilage in a mammal app. A series of primes for amplifying important regions of the gene for type II procollagen (COL2A1) by PCR and the primers for sequencing the resulting PCR fragments are given and the results summarized. The method is further improved by using denaturing gradient gel electrophoresis and primers contg. a GC-rich clamp sequence to facilitate the detection of the disease-causing mutations.

=> d fhitseq 5

L3 ANSWER 5 OF 5 CAPLUS COPYRIGHT 2003 ACS
 IT 158030-00-3
 RL: ANST (Analytical study)
 (primer, for detecting mutations on gene for type II procollagen, in
 detn. of predisposition. for osteoarthritis)
 RN 158030-00-3 CAPLUS
 CN DNA, d(C-A-T-G-G-A-G-G-A-G-T-G-A-T-A-T-T-C) (9CI) (CA INDEX NAME)

SEQ 1 catggaggag tgatattc